FOR

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [7]: C12Q 1/70, 1/68, C12N 15/10, 15/34, 1/21, C07K 14/01, C12Q 1/18 — A2

(11) International Publication Number: **WO 00/32825**

(43) International Publication Date: 8 June 2000 (08.06.00)

(21) International Application Number: PCT/IB99/02040

(22) International Filing Date: 3 December 1999 (03.12.99)

(30) Priority Data:

60/110,992	3 December 1998 (03.12.98)	US
09/326,144	3 June 1999 (03.06.99)	US
09/407,804	28 September 1999 (28.09.99)	US
60/157,218	30 September 1999 (30.09.99)	US
60/168,777	1 December 1999 (01.12.99)	US
09/454,252	2 December 1999 (02.12.99)	US

(71) Applicant *(for all designated States except US)*: PHAGETECH, INC. [CA/CA]; Place du Parc, Case Postale 387, Montreal H2W 2N9 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: PELLETIER, Jerry [CA/CA]; 8 Lakeview, Baie D'Urfe, Quebec H9X 3B1 (CA). GROS, Phillippe [CA/CA]; 107 Montrose, St. Lambert, Quebec J4R 1X4 (CA). DUBOW, Michael [CA/CA]; 4901 Coolbrook Avenue, Montreal, Quebec H3X 2K8 (CA).

(74) Agents: MORROW, Joy, D. et al.; Smart & Biggar, 900 – 55 Metcalfe Street, P.O. Box 2999, Station D, Ottawa, Ontario K1P 5Y6 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: DEVELOPMENT OF NOVEL ANTI-MICROBIAL AGENTS BASED ON BACTERIOPHAGE GENOMICS

(57) Abstract

A method for identifying suitable targets for antibacterial agents based on identifying targets of bacteriophage–encoded proteins is described. Also described are compositions useful in the identification methods and in inhibiting bacterial growth, and methods for preparing and using such compositions.

This can be done redundantly to provide the corrected sequence or to confirm that the described sequence is correct. Alternatively, a particular sequence or sequences can be identified and isolated as an insert or inserts in a phage genomic library and isolated, amplified, and sequenced by standard methods. Confirmation or correction of a nucleotide sequence for a phage gene provides an amino acid sequence of the encoded product by merely reading off the amino acid sequence according to the normal codon relationships and/or expressed in a standard expression system and the polypeptide product sequenced by standard techniques. The sequences described herein thus provide unique identification of the corresponding genes, coding sequences, and other sequences, allowing those sequences to be used in the various aspects of the present invention.

In other aspects, the invention provides recombinant vectors and cells harboring at least one of the phage ORFs or portion thereof, or bacterial target sequences described herein. As understood by those skilled in the art, vectors may be provided in different forms, including, for example, plasmids, cosmids, and virus-based vectors. See, *e.g.*, Maniatis, T. et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor University Press, Cold Spring, N.Y.; See also, Ausubel, F.M. et al. (eds.) (1994) Current Protocols in Molecular Biology. John Wiley & Sons, Secaucus, N.J.

In preferred embodiments, the vectors will be expression vectors, preferably shuttle vectors that permit cloning, replication, and expression within bacteria. An "expression vector" is one having regulatory nucleotide sequences containing transcriptional and translational regulatory information that controls expression of the nucleotide sequence in a host cell. Preferably the vector is constructed to allow amplification from vector sequences flanking an insert locus. In certain embodiments, the expression vectors may additionally or alternativley support expression, and/or replication in animal, plant and/or yeast cells due to the presence of suitable regulatory sequences, *e.g.*, promoters, enhancers, 3' stabilizing sequences, primer sequences, etc. In preferred embodiments, the promoters are inducible and specific for the system in which expression is desired, *e.g.*, bacteria, animal, plant, or yeast. The vectors may optionally encode a "tag" sequence or sequences to facilitate protein purification. Convenient restriction enzyme cloning sites and suitable selective marker(s) are also optionally included. Such selective markers can be, for example, antibiotic resistance markers or markers which supply an essential nutritive growth factor to an otherwise deficient mutant host, *e.g.*, tryptophan, histidine, or leucine in the Yeast Two-Hybrid systems described below.

The term "recombinant vector" relates to a single- or double-stranded circular nucleic acid molecule that can be transfected into cells and replicated within or independently of a cell genome. A circular double-stranded nucleic acid molecule can be cut and thereby linearized upon treatment with appropriate restriction enzymes. An assortment of nucleic acid vectors, restriction enzymes, and the knowledge of the nucleotide sequences cut by restriction enzymes are readily available to those skilled in the art. A nucleic acid molecule encoding a desired product can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together. Preferably the vector is an expression vector, *e.g.*, a shuttle expression vector as described above.

By " recombinant cell" is meant a cell possessing introduced or engineered nucleic acid sequences, *e.g.*, as described above. The sequence may be in the form of or part of a vector or may be integrated into the host cell genome. Preferably the cell is a bacterial cell.

In another aspect, the invention also provides methods for identifying and/or screening compounds "active on" at least one bacterial target of a bacteriophage inhibitor protein or RNA. Preferred embodiments involve contacting such a bacterial target or targets (e.g., bacterial target proteins) with a test compound, and determining whether the compound binds to or reduces the level of activity of the bacterial target (*e.g.*, a bacterial target protein). Preferably this is done either *in vivo* (*i.e.*, in a cell-based assay) or *in vitro*, e.g., in a cell-free system under approximately physiological conditions.

The compounds that can be used may be large or small, synthetic or natural, organic or inorganic, proteinaceous or non-proteinaceous. In preferred embodiments, the compound is a peptidomimetic, as described herein, a bacteriophage inhibitor protein or fragment or derivative thereof, preferably an "active portion", or a small molecule.

In preferred embodiments, the bacterial target is a target of a phage ORF identified herein, *e.g., S. aureus* phage 44AHJD ORF 1, 9, or 12, *Streptococcus pneumoniae* phage Dp-1 ORF 001, 002, 004, 008, 010, 013, 016, 021, 029, 030, 038, or 041, or *Enterococcus* sp. phage 182 ORF 002, 008, or 014.

In particular embodiments, the methods include the identification of bacterial targets or the site of action of an inhibitor on a bacterial target as described above or otherwise described herein.

In embodiments involving binding assays, preferably binding is to a fragment or portion of a bacterial target protein, where the fragment includes less than 90%, 80%, 70%, 60%, 50%, 40%, or 30% of an intact bacterial target protein. Preferably,

the at least one bacterial target includes a plurality of different targets of bacteriophage inhibitor proteins, preferably a plurality of different targets. The plurality of targets can be in or from a plurality of different bacteria, but preferably is from a single bacterial species.

A "method of screening" refers to a method for evaluating a relevant activity or property of a large plurality of compounds (e.g., a bacteria-inhibiting activity), rather than just one or a few compounds. For example, a method of screening can be used to conveniently test at least 100, more preferably at least 1000, still more preferably at least 10,000, and most preferably at least 100,000 different compounds, or even more.

In the context of this invention, the term "small molecule" refers to compounds having molecular mass of less than 2000 Daltons, preferably less than 1500, still more preferably less than 1000, and most preferably less than 600 Daltons. Preferably but not necessarily, a small molecule is not an oligopeptide.

In a related aspect or in preferred embodiments, the invention provides a method of screening for potential antibacterial agents by determining whether any of a plurality of compounds, preferably a plurality of small molecules, is active on at least one target of a bacteriophage inhibitor protein or RNA. Preferred embodiments include those described for the above aspect, including embodiments which involve determining whether one or more test compounds bind to or reduce the level of activity of a bacterial target, and embodiments which utilize a plurality of different targets as described above.

The identification of bacteria-inhibiting phage ORFs and their encoded products also provides a method for identifying an active portion of such an encoded product. This also provides a method for identifying a potential antibacterial agent by identifying such an active portion of a phage ORF or ORF product. In preferred embodiments, the identification of an active portion involves one or more of mutational analysis, deletion analysis, or analysis of fragments of such products. The method can also include determination of a 3-dimensional structure of an active portion, such as by analysis of crystal diffraction patterns. In further embodiments, the method involves constructing or synthesizing a peptidomimetic compound, where the structure of the peptidomimetic compound corresponds to the structure of the active portion. In this context, "corresponds" means that the peptidomimetic compound structure has sufficient similarities to the structure of the active portion that the peptidomimetic will interact with the same molecule as the phage protein and preferably will elicit at least one cellular response in common which relates to the inhibition of the cell by the phage protein.

In preferred embodiments, the ORF or ORF product is or is derived or obtained from *S. aureus* phage 44AHJD ORF 1, 9, or 12, *Streptococcus pneumoniae* phage Dp-1 ORF 001, 002, 004, 008, 010, 013, 016, 021, 029, 030, 038, or 041, or *Enterococcus* sp. phage 182 ORF 002, 008, or 014 or product thereof.

The methods for identifying or screening for compounds or agents active on a bacterial target of a phage-encoded inhibitor can also involve identification of a phage-specific site of action on the target.

Preferably in the methods for identifying or screening for compounds active on such a bacterial target, the target is uncharacterized; the target is from an uncharacterized bacterium from Table 1; the site of action is a phage-specfic site of action.

Further embodiments include the identification of inhibitor phage ORFs and bacterial targets as in aspects above.

An "active portion" as used hercin denotes an epitope, a catalytic or regulatory domain, or a fragment of a bacteriophage inhibitor protein that is responsible for, or a significant factor in, bacterial target inhibition. The active portion preferably may be removed from its contiguous sequences and, in isolation, still effect inhibition.

By "mimetic" is meant a compound structurally and functionally related to a reference compound that can be natural, synthetic, or chimeric. In terms of the present invention, a "peptidomimetic," for example, is a compound that mimics the activity-related aspects of the 3-dimensional structure of a peptide or polyeptide in a non-peptide compound, for example mimics the structure of a peptide or active portion of a phage- or bacterial ORF-encoded polypeptide.

A related aspect provides a method for inhibiting a bacterial cell by contacting the bacterial cell with a compound active on a bacterial target of a bacteriophage inhibitor protein or RNA, where the target was uncharacterized. In preferred embodiments, the compound is such a protein, or a fragment or derivative thereof; a structural mimetic, *e.g.*, a peptidomimetic, of such a protein or fragment; a small molecule; the contacting is performed *in vitro*, the contacting is performed *in vivo* in an infected or at risk organism, *e.g.*, an animal such as a mammal or bird, for example, a human, or other mammal described herein; the bacterium is selected from a genus and/or species listed in Table 1; the bacteriophage inhibitor protein is uncharacterized; the bacteriophage inhibitor protein is from an uncharacterized phage listed in Table 1; the phage inhibitor protein is from one of *S. aureus* phage 44AHJD ORF 1, 9, or 12, *Streptococcus pneumoniae* phage Dp-1 ORF 001, 002, 004, 008, 010, 013, 016, 021, 029, 030, 038, or 041, or *Enterococcus* sp. phage 182 ORF 002, 008, or 014.

In the context of targets in this invention, the term "uncharacterized" means that the target was not recognized as an appropriate target for an antibacterial agent prior to the filing of the present application or alternatively prior to the present invention. Such lack of recognition can include, for example, situations where the target and/or a nucleotide sequence encoding the target were unknown, situations where the target was known, but where it had not been identified as an appropriate target or as an essential cellular component, and situations where the target was known as essential but had not been recognized as an appropriate target due to a belief that the target would be inaccessible or otherwise that contacting the cell with a compound active on the target *in vitro* would be ineffective in cellular inhibition, or ineffective in treatment of an infection. Methods described herein utilizing bacterial targets, *e.g.*, for inhibiting bacteria or treating bacterial infections, can also utilize "uncharacterized target sites", meaning that the target has been previously recognized as an appropriate target for an antibacterial agent, but where an agent or inhibitor of the invention is used which acts at a different site than that at which the previously utilized antibacterial agent, *i.e.*, a phage-specific site. Preferably the phage-specific site has different functional characteristics from the previously utilized site. In the context of targets or target sites, the term "phage-specific" indicates that the target or site is utilized by at least one bacteriophage as an inhibitory target and is different from previously identified targets or target sites.

In the context of this invention, the term "bacteriophage inhibitor protein" refers to a protein encoded by a bacteriophage nucleic acid sequence which inhibits bacterial function in a host bacterium. Thus, it is a bacteria-inhibiting phage product.

In the context of this invention, the phrase "contacting the bacterial cell with a compound active on a bacterial target of a bacteriophage inhibitor protein" or equivalent phrases refer to contacting with an isolated, purified, or enriched compound or a composition including such a compound, but specifically does not rely on contacting the bacterial cell with an intact phage which encodes the compound. Preferably no intact phage are involved in the contacting.

Related aspects provide methods for prophylactic or therapeutic treatment of a bacterial infection by administering to an infected, challenged or at risk organism a therapeutically or prophylactically effective amount of a compound active on a target of a bacteriophage inhibitor protein or RNA, or as described for the previous aspect. Preferably the bacterium involved in the infection or risk of infection produces the identified target of the bacteriophage inhibitor protein or alternatively produces a homologous target compound. In preferred embodiments, the host organism is a plant or animal, preferably a mammal or bird, and more preferably, a human or other

```
253      M K I P A V A L M G V G G G N Q I N L L K R L P Y R N I
51801    gttctagcacttgaccctgataacgctgggcagacagcgcaggaaaaactctaccgacagttaaagcgaagcaaggtcgttaga
281      V L A L D P D N A G Q T A Q E K L Y R Q L K R S K V V R

51885    tttttgaactaccctaaagagttctatgataataagtgggatataaacgaccatccggaattattaaatttttaatgatttagtc
309      F L N Y P K E F Y D N K W D I N D H P E L L N F N D L V
51969    ttgtag 51974
337      L *
dp1ORF015
3793     atgggatttaatctatacttcgcaggaggtcacgctattagcactgacgattatttgaaggaaagaggagccaatcgcctattc
1        M G F N L Y F A G G H A I S T D D Y L K E R G A N R L F
3877     aatcaactgtacgaaagaaacgggattggcaaaaggtggattgagcataagaaaaccaatccaagcactacttcaaaactattc
29       N Q L Y E R N G I G K R W I E H K K T N P S T T S K L F
3961     gtcgactctagtgcatattctgctcataccaaaggggctgaagttgacattgacgcctatatcgaatacgtgaatgataacgtg
57       V D S S A Y S A H T K G A E V D I D A Y I E Y V N D N V
4045     ggaatgtttgactgtatcgccgaactcgataaaattcctggtgtatttagacagcctaagacacgtgaacagcttttggaagca
85       G M F D C I A E L D K I P G V F R Q P K T R E Q L L E A
4129     ccacaaatttcttgggataattatctatacatgcgcgagcgaatggttgagaaagacaagctcttacctatttttccatatggga
113      P Q I S W D N Y L Y M R E R M V E K D K L L P I F H M G
4213     gaagactttaaatggctcaacttgatgctcgaaactacattcgaaggcggaaagcatattccttacattggaatttcaccagcc
141      E D F K W L N L M L E T T F E G G K H I P Y I G I S P A
4297     aatgactcgactacgaagcataaagacaagtggatggaaagagtattcgaagttattcgaaacagttctaatccagacgttaag
169      N D S T T K H K D K W M E R V F E V I R N S S N P D V K
4381     actcacgcatttgggatgacagttactagccaattagagcgtcacccattctatagcgccgactctacttctgtactgctcaca
197      T H A F G M T V T S Q L E R H P F Y S A D S T S V L L T
4465     ggagcgatgggaaacattatgacgtcaaaaggattagttgacttgtcacagaagaatggaggaattgatgctgtccgtaggctg
225      G A M G N I M T S K G L V D L S Q K N G G I D A V R R L
4549     ccaaaaccggttcaagttgaaattgaatccattatcgaagaaactggagcgcattttagcctagagcaattagttgaggactat
253      P K P V Q V E I E S I I E E T G A H F S L E Q L V E D Y
4633     aaacttcgagcattgttcaatgttcaatacatgctgaattgggcagagaactatgaattcaagggaattaaaaatcgtcaacgt
281      K L R A L F N V Q Y M L N W A E N Y E F K G I K N R Q R
4717     cgactattttag 4728
309      R L F *
dp1ORF016
43413    atgggagtcgatattgaaaaaggcgttgcgtggatgcaggcccgaaagggtcgagtatcttatagcatggactttcgagacggt
1        M G V D I E K G V A W M Q A R K G R V S Y S M D F R D G
43497    cctgatagctatgactgctcaagttctatgtactatgctctccgctcagccggagcttcaagtgctggatgggcagtcaatact
29       P D S Y D C S S S M Y Y A L R S A G A S S A G W A V N T
43581    gagtacatgcacgcatggcttattgaaaacggttatgaactaattagtgaaaatgctccgtgggatgctaaacgaggcgacatc
57       E Y M H A W L I E N G Y E L I S E N A P W D A K R G D I
43665    ttcatctggggacgcaaaggtgctagcgcaggcgctggaggtcatacagggatgttcattgacagtgataacatcattcactgc
85       F I W G R K G A S A G A G G H T G M F I D S D N I I H C
43749    aactacgcctacgacggaatttccgtcaacgaccacgatgagcgttggtactatgcaggtcaaccttactactacgtctatcgc
113      N Y A Y D G I S V N D H D E R W Y Y A G Q P Y Y Y V Y R
43833    ttgactaacgcaaatgctcaaccggctgagaagaaacttggctggcagaaagatgctactggtttctggtacgctcgagcaaac
141      L T N A N A Q P A E K K L G W Q K D A T G F W Y A R A N
43917    ggaacttatccaaaagatgagttcgagtatatcgaagaaaacaagtcttggttctactttgacgaccaaggctacatgctcgct
169      G T Y P K D E F E Y I E E N K S W F Y F D D Q G Y M L A
44001    gagaaatggttgaaacatactgatggaaattggtattggttcgaccgtgacggatacatggctacgtcatggaaacggattggc
197      E K W L K H T D G N W Y W F D R D G Y M A T S W K R I G
44085    gagtcatggtactacttcaatcgcgatggttcaatggtaaccggttggattaagtattacgataattggtattattgtgatgct
225      E S W Y Y F N R D G S M V T G W I K Y Y D N W Y Y C D A
44169    accaacggcgacatgaaatcgaatgcgtttatccgttataacgacggctggtatctactattaccggacggacgtctggcagat
253      T N G D M K S N A F I R Y N D G W Y L L L P D G R L A D
44253    aaacctcaattcaccgtagagccggacgggctcattactgctaaagtttaa 44303
281      K P Q F T V E P D G L I T A K V *
dp1ORF017
11242    atgattggacagggacttgttaaatctaccatttcgaaatggaaacaacttccaaaatatataatcgtcgaaggtgaagtaggt
1        M I G Q G L V K S T I S K W K Q L P K Y I I V E G E V G
11326    tcaggacggaagacctaatccgttatattgcttcgaaatttgacgctgattctattgtagtaggaacgagtgtagatgacatt
29       S G R K T L I R Y I A S K F D A D S I V V G T S V D D I
11410    cgaaacatcattcaggatgcacagactatttttcaaggcgagaatctacgtgatagacggaaatagcctgtcaatgtcagctct
57       R N I I Q D A Q T I F K A R I Y Y V I D G N S L S M S A L
11494    aactcgcttttgaagatagcggaagagccacctttaaactgtcatatagccatgactgttgatagcatcaataatgctttacct
85       N S L L K I A E E P P L N C H I A M T V D S I N N A L P
11578    acgcttgcaagtagagcaaaagttctaaccatgctaccttatactaatgaagagaaaatgcagtttgtcaagtcctacaagaag
113      T L A S R A K V L T M L P Y T N E E K M Q F V K S Y K K
11662    gtagatacttcaggaattgacgaccgagcgattgtagactattgcaatcttgccagcaatcttcaaatgcttgaagacatatta
141      V D T S G I D D R A I V D Y C N L A S N L Q M-L E D I L
11746    gaatatggcgcagaagagctatttgaaaaggttacaacattttatgacttaatatgggaggcaagtgctaggaattcgctaaag
169      E Y G A E E L F E K V T T F Y D L I W E A S A S N S L K
11830    gttactaattggctcaaatttaaggaaactgatgaaggaaaaattgagcctaaacttttcctcaactgttcttttaaattggtcg
197      V T N W L K F K E T D E G K I E P K L F L N C L L N W S
11914    acagttgtcatcaggaagcactatgagaaatgtctttcgaagaacttgaggcccatgaccttttagtgagggaagcatctagg
225      T V V I R K H Y V E M S F E E L E A H D L L V R E A S R
```

```
Query: 313 PKEFYDNKWD 322
           + D  WD
Sbjct: 348 EHNYKD--WD 355



Query= sid|114837|lan|dp1ORF016 Phage dp1 ORF|43413-44303|3
       (296 letters)

>emb|CAB07986| (Z93946) N-acetylmuramoyl-L-alanine amidase [bacteriophage Dp-1]
          Length = 296

 Score =  661 bits (1686), Expect = 0.0
 Identities = 296/296 (100%), Positives = 296/296 (100%)

Query: 1   MGVDIEKGVAWMQARKGRVSYSMDFRDGPDSYDCSSSMYYALRSAGASSAGWAVNTEYMH 60
           MGVDIEKGVAWMQARKGRVSYSMDFRDGPDSYDCSSSMYYALRSAGASSAGWAVNTEYMH
Sbjct: 1   MGVDIEKGVAWMQARKGRVSYSMDFRDGPDSYDCSSSMYYALRSAGASSAGWAVNTEYMH 60

Query: 61  AWLIENGYELISENAPWDAKRGDIFIWGRKGASAGAGGHTGMFIDSDNIIHCNYAYDGIS 120
           AWLIENGYELISENAPWDAKRGDIFIWGRKGASAGAGGHTGMFIDSDNIIHCNYAYDGIS
Sbjct: 61  AWLIENGYELISENAPWDAKRGDIFIWGRKGASAGAGGHTGMFIDSDNIIHCNYAYDGIS 120

Query: 121 VNDHDERWYYAGQPYYYVYRLTNANAQPAEKKLGWQKDATGFWYARANGTYPKDEFEYIE 180
           VNDHDERWYYAGQPYYYVYRLTNANAQPAEKKLGWQKDATGFWYARANGTYPKDEFEYIE
Sbjct: 121 VNDHDERWYYAGQPYYYVYRLTNANAQPAEKKLGWQKDATGFWYARANGTYPKDEFEYIE 180

Query: 181 ENKSWFYFDDQGYMLAEKWLKHTDGNWYWFDRDGYMATSWKRIGESWYYFNRDGSMVTGW 240
           ENKSWFYFDDQGYMLAEKWLKHTDGNWYWFDRDGYMATSWKRIGESWYYFNRDGSMVTGW
Sbjct: 181 ENKSWFYFDDQGYMLAEKWLKHTDGNWYWFDRDGYMATSWKRIGESWYYFNRDGSMVTGW 240

Query: 241 IKYYDNWYYCDATNGDMKSNAFIRYNDGWYLLLPDGRLADKPQFTVEPDGLITAKV 296
           IKYYDNWYYCDATNGDMKSNAFIRYNDGWYLLLPDGRLADKPQFTVEPDGLITAKV
Sbjct: 241 IKYYDNWYYCDATNGDMKSNAFIRYNDGWYLLLPDGRLADKPQFTVEPDGLITAKV 296

Query= sid|114841|lan|dp1ORF020 Phage dp1 ORF|1864-2658|1
       (264 letters)

>emb|CAB13247| (Z99111) similar to coenzyme PQQ synthesis [Bacillus subtilis]
          Length = 243

 Score =  217 bits (548), Expect = 5e-56
 Identities = 117/248 (47%), Positives = 163/248 (65%), Gaps = 15/248 (6%)

Query: 23  MPIMEIFGPTIQGEGMVIGQKTIFIRTGGCDYHCNWCDSAFTWNGTTEPE--YITGKEAA 80
           +P++EIFGPTIQGEGMVIGQKT+F+RT GCDY C+WCDSAFTW+G+ + +  ++T +E
Sbjct: 5   IPVLEIFGPTIQGEGMVIGQKTMFVRTAGCDYSCSWCDSAFTWDGSAKKDIRWMTAEEIF 64

Query: 81  SRILKLAFNDKGEQICNHVTLTGGNPALINEPMAKMISILKEHGFKFGLETQGTRFQEWF 140
           + +       D G     +HVT++GGNPAL+ + +    I +LKE+  +  LETQGT +Q+WF
Sbjct: 65  AEL-----KDIGGDAFSHVTISGGNPALLKQ-LDAFIELLKENNIRAALETQGTVYQDWF 118

Query: 141 KEVSDITISPKPPSSGMRTNMKILEAIVDRM--NDENLDWSFKIVIFDENDLAYARDMFK 198
             + D+TISPKPPSS M TN + L+ I+  +  ND      S K+VIF++ DL +A+ + K
Sbjct: 119 TLIDDLTISPKPPSSKMVTNFQKLDHILTSLQENDRQHAVSLKVVIFNDEDLEFAKTVHK 178

Query: 199 TFEGKLRPVNYLSVGNANAY--EEGKISDRLLEKLGWLWDKVYEDPAFNNVRPLPQLHTL 256
            + G      YL VGN + +  ++  +   LL K   L DKV  D   N VR LPQLHTL
Sbjct: 179 RYPG---IPFYLQVGNDDVHTTDDQSLIAHLLGKYEALVDKVAVDAELNLVRVLPQLHTL 235

Query: 257 VYDNKRGV 264
           ++ NKRGV
Sbjct: 236 LWGNKRGV 243
```